

POLYNOMIAL APPROXIMATION OF SELF-SIMILAR MEASURES AND THE SPECTRUM OF THE TRANSFER OPERATOR

CHRISTOPH BANDT AND HELENA PEÑA

ABSTRACT. We consider self-similar measures on \mathbb{R} . The Hutchinson operator H acts on measures and is the dual of the transfer operator T which acts on continuous functions. We determine polynomial eigenfunctions of T . As a consequence, we obtain eigenvalues of H and natural polynomial approximations of the self-similar measure. Bernoulli convolutions are studied as an example.

1. INTRODUCTION

We consider self-similar measures ν on \mathbb{R} , given by affine maps $f_i(x) = t_i x + v_i$ with $0 < t_i < 1$, and probabilities p_i for $i = 1, \dots, m$. A simple example are Bernoulli convolutions where we have two mappings on an interval, $t_1 = t_2 = t$ and $p_1 = p_2 = \frac{1}{2}$. The measure ν has compact support X which is the self-similar set defined by the f_i . It is the unique probability measure which is fixed by the Hutchinson operator

$$H\mu(B) = \sum_{i=1}^m p_i \mu(f_i^{-1}(B)) \quad \text{for Borel sets } B \subset X. \quad (1)$$

H acts on the space of all finite Borel measures on X . Hutchinson [10] proved that for any initial probability measure μ_0 the sequence $\mu_n = H\mu_{n-1}$, $n = 1, 2, \dots$ converges to ν geometrically, with factor $t = \max t_i$, with respect to a certain metric. Thus 1 is an eigenvalue of H , with eigenvector ν , and there should be no other eigenvalues of modulus larger than t . The situation is a bit more complicated and will be discussed in Section 2.

For every positive integer k , there are examples where ν has a density which is k times differentiable. All Bernoulli convolutions with sufficiently large parameter $t < 1$ outside an exceptional set of dimension zero, belong to this class [12]. In this case, the first k derivatives of the density are eigenfunctions of H corresponding to the eigenvalues t, t^2, \dots, t^k . Numerical experiments indicate that this holds if the density has only $k - 1$ derivatives, and these values are the leading eigenvalues of the operator (Figure 1). We shall give an explanation for this behavior.

Hutchinson's theorem has led to well-known iteration algorithms which generate pictures of the self-similar measure ν . These methods are fast and often give a nice impression of fractal structures. In cases where ν has a density, however, they are not too accurate. It is hard to decide from such approximations whether the function is smooth or monotone since

This work was supported by Deutsche Forschungsgemeinschaft grant Ba 1332/11-1.

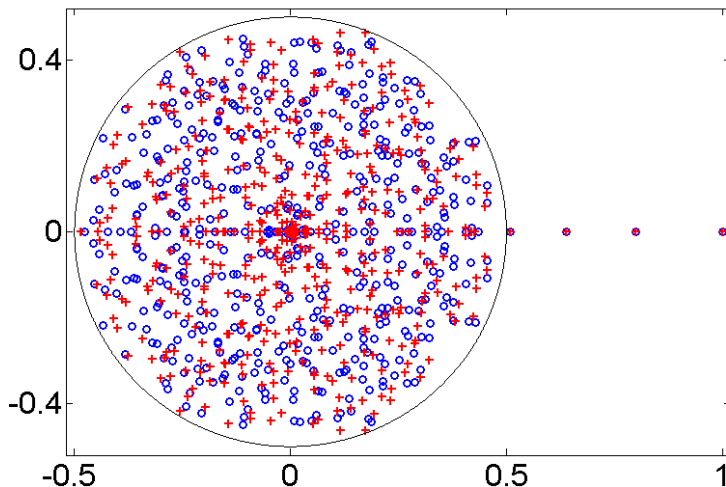


FIGURE 1. Eigenvalues of two matrix approximations of the Hutchinson operator for the Bernoulli convolution with $t = 0.8$. The matrix size, $N = 499$ and 500 , does not influence the leading eigenvalues $t^k, k = 0, \dots, 3$ while virtually all remaining eigenvalues are different. The circle has radius $\frac{1}{2}$.

one cannot distinguish genuine fractal structure and noise. Moreover, an approximation given by an iteration algorithm is just a set of data, not a mathematical object. We shall present an approximation by polynomials as an alternative. See Figure 2.

The key to our study is the *transfer operator* T acting on continuous functions $h \in C(X)$ as follows.

$$Th(x) = \sum_{i=1}^m p_i \cdot h(f_i(x)) \quad \text{for } x \in X. \quad (2)$$

The Hutchinson operator is the dual operator of T and so their spectra are equal, see for instance [11, Theorem 6.22]. Our first result says that T has polynomial eigenfunctions which are easy to determine.

Theorem 1. (*Polynomial eigenfunctions of the transfer operator, real case*)

The transfer operator T has eigenvalues $\lambda_n = \sum_{i=1}^m p_i t_i^n$ and corresponding polynomial eigenfunctions q_n of degree n for $n = 0, 1, 2, \dots$. If the mappings f_i have equal contraction factors $t_i = t$, these eigenvalues are $\lambda_n = t^n$ for $n = 0, 1, 2, \dots$.

Different versions of Theorem 1 will be discussed in Section 3. The proof provides a construction of the q_n . Then we turn to the calculation of polynomial approximations of the self-similar measure ν . In the case that the support X of ν is an interval, the following theorem holds with the usual notation

$$\langle f, g \rangle = \int_X f(x)g(x)dx \quad \text{and} \quad \nu(f) = \int_X f(x)d\nu(x) \quad \text{for } f, g \in C(X). \quad (3)$$

Theorem 2. (*Polynomial approximation of the self-similar measure*)

Given the polynomial eigenfunctions $q_0 = 1, q_1, q_2, \dots$ of the transfer operator T , let v_n denote the polynomial on X of degree n which satisfies the equations

$$\langle v_n, q_0 \rangle = 1, \quad \langle v_n, q_k \rangle = 0 \quad \text{for } k = 1, \dots, n. \quad (4)$$

Considered as density function, v_n is the best approximation of the measure ν among polynomials of degree at most n in the sense that

$$\frac{\nu(v_n)}{\|v_n\|_2} > \frac{\nu(q)}{\|q\|_2}$$

for all polynomials q of degree $\leq n$ which are not multiples of v_n . If ν has an L^2 density v then v_n converges to v in the L^2 norm.

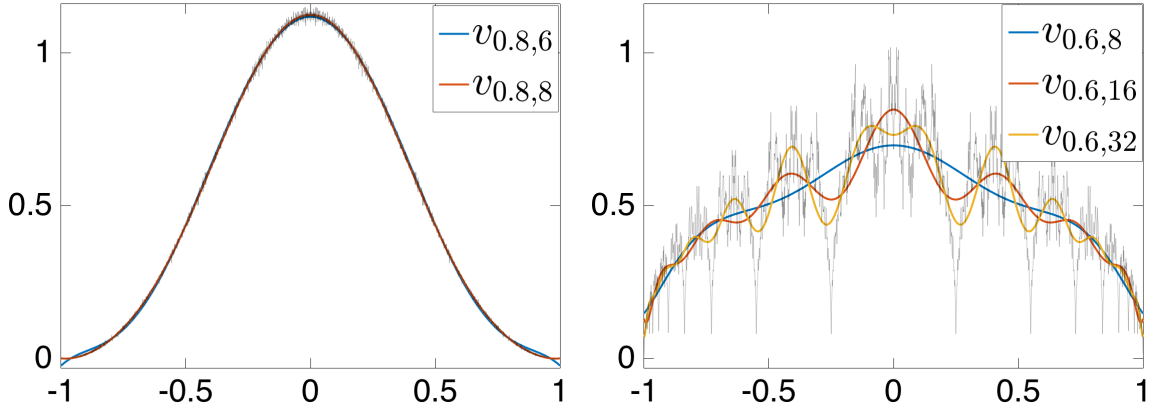


FIGURE 2. Polynomial approximations $v_{t,n}$ of the Bernoulli convolution measure ν_t in the smooth case $t = 0.8$ (left) and the more fractal case $t = 0.6$ (right). In the left picture, one could not distinguish at this scale between the approximations of degree $n \geq 8$. The gray line shows a histogram with 2000 bins based on 2^{20} points generated by the 'chaos game' algorithm.

Figure 2 shows approximations of a smooth and a more fractal self-similar measure. In Section 4 we prove Theorem 2 and discuss modifications. In Section 5 we show that the problem of finding the approximating polynomials v_n can be regarded as a moment problem, as considered in [4, 1]. As a consequence, construction of the approximation v_n amounts to solving a system of linear equations with a Hilbert matrix of coefficients. Section 6 deals with the case of Bernoulli convolutions, and Theorem 5 will give an explicit analytical representation of the q_n for this case. In the final section we indicate how singular measures can be studied with this approach.

2. EXPERIMENTS WITH THE SPECTRUM OF HUTCHINSON'S OPERATOR

Before we go into proofs, let us discuss some experiments which have motivated this work. All Bernoulli convolutions will be considered on the interval $I = [-1, 1]$. The parameter t varies between 0.5 and 1, the figures refer to $t = 0.8$. The mappings are

$$f_1(x) = tx - (1 - t) \quad \text{and} \quad f_2(x) = tx + (1 - t). \quad (5)$$

One way to think about self-similar measures is to consider ν as stationary distribution with respect to random iteration of the functions, popularized as 'chaos game'. We have a discrete time Markov process (X_0, X_1, \dots) with kernel

$$p(x, B) = \mathbb{P}(X_{n+1} \in B | X_n = x) = \frac{1}{2} \delta_B(f_1(x)) + \frac{1}{2} \delta_B(f_2(x)) \quad (6)$$

To approximate by a Markov chain (Z_0, Z_1, \dots) with N states, we divide I into N disjoint intervals I_k and introduce the transition probabilities

$$p_{kj} = \mathbb{P}(Z_{n+1} \in I_j | Z_n \in I_k) = \frac{1}{2} \frac{|f_1^{-1}(I_j) \cap I_k|}{|I_k|} + \frac{1}{2} \frac{|f_2^{-1}(I_j) \cap I_k|}{|I_k|}$$

where $|J|$ denotes the length of the interval J . The matrix $T_N = (p_{kj})_{k,j=1,\dots,N}$ represents the transfer operator T when multiplied by column vectors $(v_1, \dots, v_N)'$ which describe functions, and the Hutchinson operator H when multiplied from the left by row vectors (u_1, \dots, u_N) which describe measures. This matrix is always irreducible and aperiodic. (This follows from $p_{11} > 0$, $p_{NN} > 0$, and from the fact that an interval J is mapped either by f_1^{-1} or by f_2^{-1} to an interval which is $1/t$ times longer, unless J contains the points t and $1 - t$ and thus will be mapped to state 1 and N .)

So there is a unique left eigenvector of T_N for the eigenvalue 1 which represents the self-similar measure, and a corresponding right eigenvector which is constant since we have a Markov matrix. The convergence is fast: even if we take $N = 10^6$ and start with uniform distribution, 60 iterations yield sufficient accuracy. Calculations were performed with MATLAB. To determine eigenvalues and eigenvectors, we took N between 400 and 2000. For each parameter t , different N were tested, with partitions into intervals of equal lengths and randomly perturbed partitions.

The pattern of eigenvalues was always the same, as shown in Figure 1. The values $1, t, t^2, \dots, t^k$ were eigenvalues, as long as $t^k > \frac{1}{2}$. The remaining eigenvalues were spread inside the circle $|\lambda| < \frac{1}{2}$. In a few experiments, one or two percent of the points were slightly outside that circle. The distribution inside the circle could be more or less uniform, it could be more a ring, especially when t is near $\frac{1}{2}$, or zero could be a multiple eigenvalue, but these properties usually changed with N , as indicated in Figure 1.

The leading eigenvalues $t^k > \frac{1}{2}, k = 0, 1, \dots$ did never change, and their left eigenvectors were also very stable. If a highly oscillating function is drawn for very different N , there are natural differences, but for the smooth eigenfunctions deviations were very small. In Figure 3, the function which represents the measure ν looks like a normal distribution, and this is true whenever $t > 0.8$ (cf. [14]).

In Figure 3, the eigenvector e_1 represents the negative of the derivative of e_0 . The third eigenvector seems irregular, but when it is integrated (just taking cumulative sums) we get

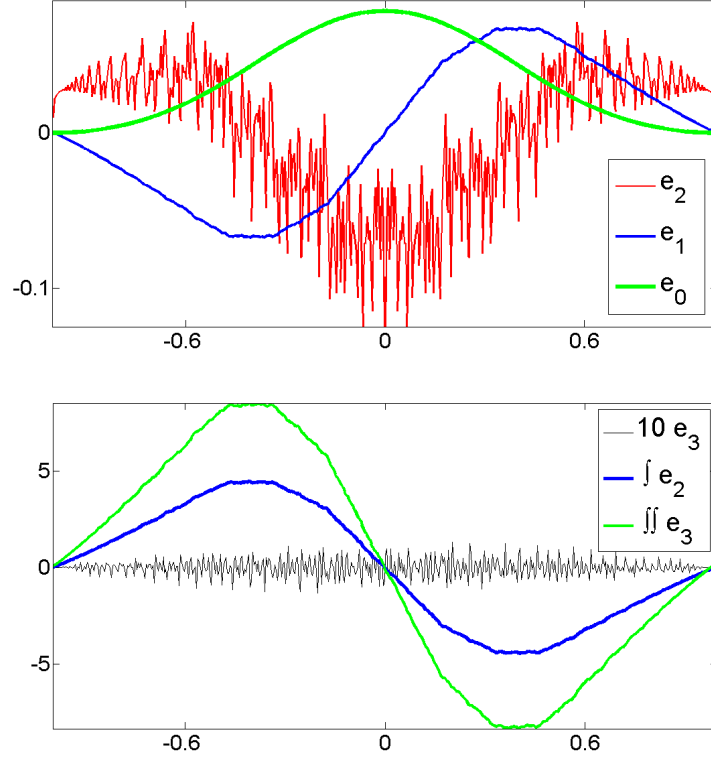


FIGURE 3. Above: leading left eigenvectors e_0, e_1, e_2 of T_N for the Bernoulli convolution with $t = 0.8$ and $N = 500$. Below: scaled eigenvector e_3 , integral of e_2 and iterated integral of e_3 . The latter two coincide with e_1 , up to a constant. Thus e_1, e_2, e_3 are the first 3 derivatives of the density e_0 of the self-similar measure.

e_1 , multiplied by a constant. This indicates that we have the second derivative of ν . And even e_3 , two times integrated, gives the same result, up to a constant. Thus we have a third derivative of ν , in spite of the fact that the second derivative seems a non-differentiable function.

For the case that ν has a k times differentiable density $\rho(x)$ it is easy to check that the derivatives are eigenvectors. The self-similarity condition (1) can be reformulated as

$$\rho(x) = \frac{\beta}{2} [\rho(\beta x - 1 + \beta) + \rho(\beta x + 1 - \beta)] =: (H\rho)(x) \quad \text{with } \beta = \frac{1}{t}.$$

Taking derivatives we get $\rho' = \beta \cdot H\rho'$, that is, $H\rho' = t\rho'$, and $H\rho'' = t^2 \cdot \rho''$ etc. We cannot prove, however, that there are no other eigenfunctions with $|\lambda| > \frac{1}{2}$. We also do not explain the fact that we get one more derivative than expected. As shown in Figure 4, this even holds when the density function is not at all differentiable. For the non-fractal case $t = \frac{1}{2}$

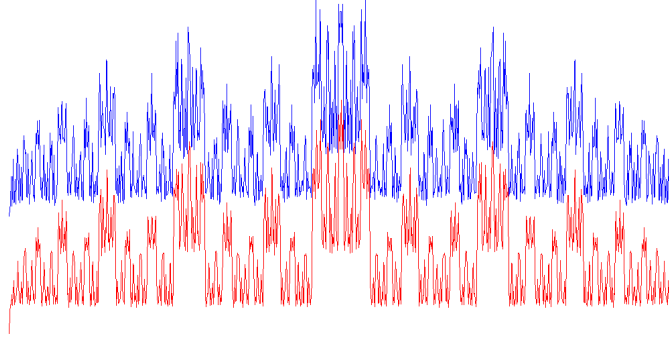


FIGURE 4. $\beta = 1/t = 1.84$ was chosen near the Pisot parameter 1.8393... Below: the self-similar measure. Above: cumulative sum of second eigenvector. Although ν is not a differentiable function, the second eigenvector looks like a derivative of ν .

this can be checked: ν has density equal $\frac{1}{2}$, and the distributional derivative $\frac{1}{2}(\delta_{-1} - \delta_1)$ is an eigenfunction of H with eigenvalue $\frac{1}{2}$.

Pisot parameters play a large role in Bernoulli convolutions since Erdős proved 1939 that ν has no density in this case [15, 14]. In our numerical study, they did not look very special, although the pattern of eigenvalues seemed somewhat different, and for the golden mean a discontinuous eigenfunction with $\lambda = -\frac{1}{2}$ was identified. Figure 4, for example, is very near to the so-called Tribonacci parameter, and still we have the 'derivative' as second eigenvector.

All eigenvalues with $|\lambda| < \frac{1}{2}$ have extremely noisy left eigenvectors, although they are symmetric, odd or even functions. Since they change with N and with slight changes of partition, they can be interpreted as noise. This is highly correlated noise, however. If these functions are integrated two times, they still have a fractal appearance. If white noise is integrated two times, a smooth function will result.

The right eigenvectors of T_N look much more regular. Those which correspond to $\lambda = t^k > \frac{1}{2}$ are polynomials, as shown in Figure 5. This will be proved below. The other right eigenvectors represent noise. They are all odd or even, and rather smooth functions – more smooth than left eigenvectors after two integrations. The duality between polynomials and special functions somewhat resembles the classical Sturm-Liouville operators which lead to special functions and associated polynomials of Legendre, Hermite etc, although our operator T is not self-adjoint and we have only a finite set of 'special functions'.

What about the circle $|\lambda| = \frac{1}{2}$? For $t = \frac{1}{2}$ there is an explanation. In this case $Th(x) = \frac{1}{2}(h(\frac{x-1}{2}) + h(\frac{x+1}{2}))$ and T has all points λ of the unit circle as eigenvalues. An eigenfunction for λ is $h(x) = \sum_{n=0}^{\infty} \lambda^n \exp(\pi k 2^n i x)$ where k is a fixed odd number. See Driebe [7] and the literature quoted there. $h(x)$ is a Weierstrass function. It is well-known

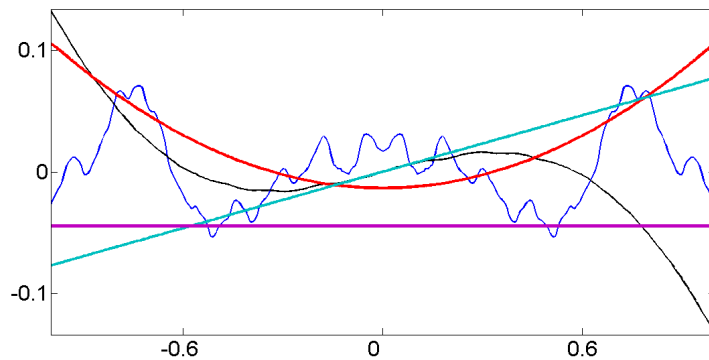


FIGURE 5. Right eigenvectors are polynomials for $\lambda > \frac{1}{2}$. First eigenvector with $|\lambda| < \frac{1}{2}$ shown for comparison. Parameters as in Figure 3.

that $h(x)$ is differentiable only for $|\lambda| < \frac{1}{2}$, see [3] and its references. Apparently, numerical experiments only detect eigenvalues of differentiable functions. This point seems worth a further study.

3. EIGENPOLYNOMIALS OF THE TRANSFER OPERATOR

We now prove Theorem 1, which concerns the existence of polynomial eigenfunctions of the transfer operator for affine IFS on \mathbb{R} . We consider affine maps $f_i(x) = t_i x + v_i$ on \mathbb{R} with $0 < t_i < 1$ for $i = 1, \dots, m$. The self-similar set X given by the f_i is the unique compact set which fulfills the equation $X = \bigcup_{i=1}^m f_i(X)$. Moreover, we have probabilities $p_i > 0$ with $\sum_{i=1}^m p_i = 1$, and the self-similar measure ν on X is defined by equation (1). The only assumption which we need is that X consists of more than one point. Then X has infinitely many points, and each polynomial on \mathbb{R} is determined by its values on X .

Proof of Theorem 1. The transfer operator T defined in (2) with mappings $f_i(x) = t_i x + v_i$ for $i = 1, \dots, m$ maps the polynomial $b_k(x) = x^k$ to

$$Tb_k(x) = \sum_{i=1}^m p_i (t_i x + v_i)^k = \sum_{l=0}^k \sum_{i=1}^m p_i \binom{k}{l} t_i^l v_i^{k-l} x^l \quad (7)$$

which is a linear combination of $b_0(x), b_1(x), \dots, b_k(x)$ and is thus a polynomial of degree k . This shows that the space of polynomials of degree $\leq k$ is an invariant subspace of T . The action of T on the polynomials of degree $\leq n$ is then represented, with respect to the basis polynomials $b_0, b_1, b_2, \dots, b_n$, by the matrix

$$T^{(n)} = \begin{pmatrix} 1 & * & * & \dots & * \\ 0 & \sum_{i=1}^m p_i t_i & * & \dots & * \\ 0 & 0 & \sum_{i=1}^m p_i t_i^2 & \dots & * \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \sum_{i=1}^m p_i t_i^n \end{pmatrix}$$

with entries

$$T_{lk}^{(n)} = \binom{k}{l} \sum_{i=1}^m p_i t_i^l v_i^{k-l} \quad \text{for } 0 \leq l \leq k \leq n \quad (8)$$

and 0 else. Note that $Tb_k = \sum_{l=0}^k T_{lk}^{(n)} b_l$. The invariant subspaces show up in the upper-triangular shape. The eigenvalues of $T^{(n)}$ are the diagonal entries $\lambda_k = \sum_{i=1}^m p_i t_i^k$, $k = 0, 1, \dots, n$. Since $0 < t_i < 1$, they satisfy $\lambda_0 = 1 > \lambda_1 > \dots > \lambda_n$. All eigenvalues are different and the right eigenvectors are polynomial eigenfunctions $q_0 = 1, q_1, \dots, q_n$ of T of degree $0, 1, \dots, n$ respectively. \square

Theorem 1 can be extended to some other IFS.

Remark 1. (Complex case)

Consider an IFS on \mathbb{C} with mappings $f_i(z) = t_i z + v_i$ with $t_i, v_i \in \mathbb{C}$, $i = 1, \dots, m$, and probabilities p_1, \dots, p_m . In this case the self-similar set X is a subset of \mathbb{C} . The operator T defined in equation (2) acts on the complex vector space $C(X)$ of continuous functions on X and has eigenfunctions and eigenvalues as in the real case. But now we have complex polynomials, and we will add their conjugates. The proof of Theorem 1 goes through when we regard the $b_k(z) = z^k$ as complex polynomials. With the additional assumption that the complex eigenvalues $\lambda_0 = 1, \lambda_1, \lambda_2, \dots$ are all different, we see that T has complex eigenpolynomials $q_0 = 1, q_1, q_2, \dots$. The 'conjugate polynomials' $1, \bar{z}, \bar{z}^2, \dots$ yield other invariant subspaces of T . Namely, \bar{z}^k is mapped into a linear combination of $1, \bar{z}, \dots, \bar{z}^k$. The action of T on $1, \bar{z}, \dots, \bar{z}^n$ is then represented by the matrix $\overline{T^{(n)}}$ which yields the conjugate eigenpolynomials $\bar{q}_1, \bar{q}_2, \dots$ for the eigenvalues $\bar{\lambda}_1, \bar{\lambda}_2, \dots$.

Remark 2. (d-dimensional case)

On \mathbb{R}^d the polynomials $q(x) = q(x_1, \dots, x_d)$ of degree n are linear combinations of the monomials $b_k(x) = x_1^{k_1} x_2^{k_2} \dots x_d^{k_d}$ where $k = (k_1, \dots, k_d)$ runs through the vectors of nonnegative integers with $k_1 + \dots + k_d = n$. We consider an affine IFS $f_i(x) = L_i x + v_i$, $i = 1, \dots, m$ where each L_i is a $d \times d$ matrix. Then it is easy to check that the space of polynomials of degree $\leq n$ is an invariant subspace of the transfer operator T for each n .

To get a triangular matrix $T^{(n)}$, one has to order the basis vectors b_k as above with respect to ascending degree $n = 0, 1, 2, \dots$. Moreover, one has to impose conditions on the maps. We assume diagonal matrices, as used for Bedford-McMullen carpets and Gatzouras-Lalley carpets: L_i has diagonal vector $\ell_i = (\ell_{i1}, \dots, \ell_{id})$ and zeros outside the diagonal. Then

$b_k(L_i x) = b_k(\ell_i) b_k(x)$, and equation (7) obtains the form

$$Tb_k(x) = \sum_{i=1}^m p_i b_k(L_i x + v_i) = \sum_{i=1}^m p_i b_k(\ell_i) b_k(x) + \text{terms of degree smaller than } n.$$

So we get an upper triangular matrix $T^{(n)}$ with diagonal entries $\lambda_k = \sum_{i=1}^m p_i b_k(\ell_i)$ which are the eigenvalues. If all f_i have the same matrix L with diagonal ℓ , then $\lambda_k = b_k(\ell)$. In case that the λ_k are all different, the eigenpolynomials q_k form a basis for the space of polynomials of degree $\leq n$. A similar remark holds if the L_i are upper triangular matrices.

4. APPROXIMATION OF THE SELF-SIMILAR MEASURE

We proved that the $(n+1)$ -dimensional vector space P_n of polynomials of degree at most n is invariant under T and contains n linearly independent non-constant eigenpolynomials. These eigenpolynomials are all orthogonal to the self-similar measure ν in the sense that

$$\nu(q_k) = \int_X q_k(x) d\nu(x) = 0 \quad \text{for } k = 1, 2, \dots \quad (9)$$

This follows from the duality of the operators T and H , and the fact that the eigenvalue of q_k fulfils $\lambda_k \neq 1$ for $k \geq 1$ by a standard argument:

$$\nu(q_k) = (H\nu)(q_k) = \nu(Tq_k) = \nu(\lambda_k q_k) = \lambda_k \nu(q_k).$$

So there is one dimension left in the space, and Theorem 2 says that the vector v_n which is orthogonal to all eigenpolynomials, considered as a density function, is the best possible approximation of ν on P_n . Moreover, the space of all polynomials is dense in $C(X)$, so that the sequence of density functions v_n is in some respect the best approximation of ν by continuous density functions. We now prove Theorem 2, assuming that the self-similar set X is an interval. We use the scalar product (3) and the norm $\|q\| = \sqrt{\langle q, q \rangle}$.

Proof of Theorem 2. Let H be the n -dimensional hyperplane generated by the eigenpolynomials q_1, \dots, q_n in the $(n+1)$ -dimensional space P_n of polynomials of degree $\leq n$. For $h \in H$ we have $\nu(h) = 0$ and $\langle v_n, h \rangle = 0$. A polynomial $q \in P_n \setminus H$ can be written as $q = \alpha v_n + h$ with $\alpha \neq 0$ and $h \in H$. This implies

$$\nu(q) = \alpha \nu(v_n) \quad \text{and} \quad \langle v_n, q \rangle = \alpha \langle v_n, v_n \rangle.$$

We see that $\nu(q)/\langle v_n, q \rangle = \nu(v_n)/\langle v_n, v_n \rangle$ is a constant. The assumption $\nu(1) = \langle v_n, 1 \rangle = 1$ shows that this constant equals one, and that $\nu(v_n) > 0$. The equality $\nu(q) = \langle v_n, q \rangle$ says that v_n is the *generating vector* of the linear form ν on the Hilbert space P_n . The Cauchy-Schwartz inequality $\nu(q) \leq \|q\| \cdot \|v_n\|$ directly implies the inequality of Theorem 2.

Now we show L^2 convergence of the v_n , for the case that ν has an L^2 density v . In this case the above proof implies that v_n is the orthogonal projection from v onto the space P_n . Thus $\|v_n - v\|$ equals the distance from v to P_n which converges to zero by the Weierstrass theorem and the fact that continuous functions are dense in L^2 , see for example [16, Theorem 1.10.18, Proposition 1.10.4]. \square

The case that ν does not possess an L^2 density will be discussed in Section 7. L^2 convergence applies in particular to the Bernoulli convolutions treated in Section 6 since their self-similar measures belong to L^2 for almost every parameter, cf. [15, 14]. We assumed that X is an interval but Theorem 2 also holds under different assumptions.

Remark 3. (IFS with open set condition)

If f_1, \dots, f_m satisfies the open set condition, it is known that the s -dimensional Hausdorff measure \mathcal{H}^s on X with $\sum_{i=1}^m t_i^s = 1$ is positive and finite [10, 8]. We can replace the Lebesgue measure in the above proof by \mathcal{H}^s , in particular in the definition (3) of $\langle f, g \rangle$. Using the Stone-Weierstrass theorem for compact Hausdorff spaces X (see e.g. [16, Theorem 1.10.18]), we obtain Theorem 2.

Remark 4. (d-dimensional case)

Based on Remark 2, Theorem 2 can be extended to the multidimensional case which is work in progress. For complex maps, as mentioned in Remark 1, there is the problem that the algebra generated by complex polynomials and their conjugates, needed for the Stone-Weierstrass theorem, is much larger than the generated vector space.

5. MOMENTS AND CONSTRUCTION OF THE BEST APPROXIMATION

The j -th moment of the probability measure ν is the integral $m_j = \nu(b_j) = \int_X x^j d\nu$, that is, the value of ν on the j -th basis vector in the polynomials, for $j = 0, 1, 2, \dots$. Characterization of measures by moments is a classical topic [13], for self-similar measures see [4, 1]. The eigenpolynomials of T provide a recursive method to compute moments of ν .

Proposition 3. (*Eigenpolynomials and moments of the self-similar measure*)

Let $q_n(x) = a_{n,0} + a_{n,1}x + \dots + a_{n,n-1}x^{n-1} + x^n$ be the eigenpolynomials of the transfer operator with leading coefficient one. The moments $m_n = \int_X x^n d\nu$, $n = 1, 2, \dots$, can be computed by

$$m_n = -a_{n,0} - a_{n,1}m_1 - \dots - a_{n,n-1}m_{n-1}$$

Proof. From $\int_X q_n d\nu = 0$ for $n = 1, 2, \dots$ we conclude

$$a_{n,0} \int_X d\nu + a_{n,1} \int_X x d\nu + \dots + \int_X x^n d\nu = a_{n,0} + a_{n,1}m_1 + \dots + m_n = 0. \quad \square$$

Now we have the moments, and we show how to compute the approximating polynomials v_n by a system of linear equations with a well-known coefficient matrix.

Proposition 4. (*System of linear equations for the approximating polynomials*)

The orthogonality relation (4) which defines the polynomial v_n is equivalent to the condition

$$\langle v_n, b_k \rangle = m_k \quad \text{for } k = 0, 1, \dots, n. \quad (10)$$

Consequently, the coefficient vector $u = (u_0, u_1, \dots, u_n)'$ of $v_n(x) = u_0 + u_1x + \dots + u_nx^n$ satisfies the linear system of equations

$$Gu = m$$

where $m = (m_0, m_1, \dots, m_n)'$ and G is the Hilbert matrix with entries

$$G_{ij} = \int_X x^{i+j} dx, \quad i, j = 0, 1, \dots, n.$$

Proof. The first assertion can be derived recursively from Proposition 3, or from the following argument. Because of (9), condition (4) says that v_n and ν generate identical linear forms on the space of polynomials generated by $q_0 = 1, q_1, \dots, q_n$. Equation (10) expresses the same fact, referring to the standard base b_0, \dots, b_n . So the two conditions are equivalent.

The explicit form of the equations (10) is easily determined:

$$\langle v_n, b_j \rangle = \int_X (u_0 + u_1 x + \dots + u_n x^n) x^j dx = u_0 \int_X x^j dx + u_1 \int_X x^{j+1} dx + \dots + u_n \int_X x^{n+j} dx$$

This is the j -th term of Gu . □

Of course, given the q_k , condition (4) can directly be considered as a system of equations. The above reformulation establishes a connection with classical matrices [9] which have been thoroughly studied by many authors, mainly for their notoriously bad condition number, see e.g. [6], [5, Example 3.3]. Moreover, the system $Gu = m$ can be solved without determining the eigenfunctions q_k when moments are directly calculated in a recursive way.

6. BERNOULLI CONVOLUTIONS

The IFS for the Bernoulli convolutions with parameter $t \in (0, 1)$ consists of the mappings $f_1(x) = tx + v_1$ and $f_2(x) = tx + v_2$ with $v_1 \neq v_2$ and probabilities $p_1 = p_2 = \frac{1}{2}$. Since there is only one conjugacy class for each t we can choose v_1 and v_2 arbitrarily. We work with the mappings

$$f_1(x) = tx - 1 + t, \quad f_2(x) = tx + 1 - t. \quad (11)$$

We denote the corresponding self-similar measure by ν_t . In case $t \in (0, \frac{1}{2})$ the measure ν_t is supported on a Cantor set and is necessarily singular. For all $t \in [\frac{1}{2}, 1)$ the support is the same interval $X = [-1, 1]$, which was a reason for choosing the mappings (11). In this case a density for ν_t might exist, and will exist for almost all $t > \frac{1}{2}$. The question for which t exactly we have a density has been studied by many authors over more than 75 years. See the surveys of Peres, Schlag, and Solomyak [15], Solomyak [14], and the recent work of Shmerkin [12].

We will now prove a formula for the eigenpolynomials of the transfer operator for the Bernoulli convolutions. Then we compute the moments of ν_t with Proposition 3. Finally, we will use these moments to compute approximations of ν_t with Proposition 4.

Theorem 5. (*Explicit formula for Bernoulli convolutions*)

For the Bernoulli convolution with mappings $f_1(x) = tx - 1 + t$ and $f_2(x) = tx + 1 - t$ the polynomial eigenfunctions q_n of the transfer operator correspond to the eigenvalues $\lambda_n = t^n$ and can be written as

$$q_n(x) = \sum_{k=0}^n a_{n,k} x^k \quad n = 0, 1, 2, \dots \quad (12)$$

where $a_{n,n} = 1$, $a_{n,n-1} = 0$, and the other coefficients are determined recursively by

$$a_{n,k} = \frac{1}{t^{n-k} - 1} \sum_{s=1}^{\lfloor \frac{n-k}{2} \rfloor} \binom{k+2s}{k} (1-t)^{2s} a_{n,k+2s} \quad k = 0, 1, \dots, n-2.$$

Thus q_n is an even function for even n and an odd function for odd n .

Proof. The action of the transfer operator on polynomials of degree n is represented by the matrix $T^{(n)}$ defined in the proof of Theorem 1. Substituting the Bernoulli convolution parameters $t_1 = t_2 = t$, $p_1 = p_2 = \frac{1}{2}$ and $v_1 = -1 + t$, $v_2 = 1 - t$ in (8) yields the entries $T_{lk}^{(n)} = \binom{k}{l} (\frac{1}{2} t^l (-1+t)^{k-l} + \frac{1}{2} t^l (1-t)^{k-l})$ for $l \leq k \leq n$ and 0 else. Non-zero entries appear only if $k-l$ is non-negative and even, in which case $T_{lk}^{(n)} = \binom{k}{l} t^l (1-t)^{k-l}$:

$$T^{(n)} = \begin{pmatrix} 1 & 0 & \binom{2}{0}(1-t)^2 & 0 & \binom{4}{0}(1-t)^4 & 0 & \dots \\ 0 & t & 0 & \binom{3}{1}t(1-t)^2 & 0 & \binom{5}{1}t(1-t)^4 & \dots \\ & 0 & t^2 & 0 & \binom{4}{2}t^2(1-t)^2 & 0 & \dots \\ \vdots & & 0 & t^3 & 0 & \binom{5}{3}t^3(1-t)^2 & \dots \\ & & & 0 & t^4 & 0 & \dots \\ & & & & \ddots & \ddots & \\ 0 & & & & & 0 & t^n \end{pmatrix}.$$

The diagonal entries $\lambda_0 = 1, \lambda_1 = t, \dots, \lambda_n = t^n$ are eigenvalues of $T^{(n)}$ and thus of T too. Let $q_n(x) = a_{n,0} + a_{n,1}x + a_{n,2}x^2 + \dots + a_{n,n-1}x^{n-1} + a_{n,n}x^n$ with $a_{n,n} = 1$ be the eigenpolynomial corresponding to the eigenvalue $\lambda_n = t^n$. The eigenvalue equation $Tp_n = t^n p_n$ can be rewritten as

$$T^{(n)}a = t^n a \quad (13)$$

in terms of the coefficient vector $a = (a_{n,0}, a_{n,1}, a_{n,2}, \dots, a_{n,n-1}, 1)'$. Suppose $a_{n,k+1}, \dots, a_{n,n}$ are known for some $k \in \{0, 1, \dots, n-1\}$. Equating the k -th coordinate in (13) yields

$$a_{n,k}t^k + \sum_{s=1}^{\lfloor \frac{n-k}{2} \rfloor} \binom{k+2s}{k} t^k (1-t)^{2s} a_{n,k+2s} = t^n a_{n,k}.$$

So we have the recursion $a_{n,k} = \frac{1}{t^{n-k} - 1} \sum_{s=1}^{\lfloor \frac{n-k}{2} \rfloor} \binom{k+2s}{k} (1-t)^{2s} a_{n,k+2s}$. Setting $k = n-1$ shows that $a_{n,n-1} = 0$. Thus q_n is a polynomial of degree n with either odd or even powers of t . \square

With the aid of the eigenpolynomials we compute the moments of ν_t , as shown in Proposition 3. Then we turn to the computation of the approximations

$$v_{t,n}(x) = u_0 + u_1x + \dots + u_nx^n$$

defined in Theorem 2. Let $m = (m_0, m_1, \dots, m_n)'$ be the first $n + 1$ moments of ν_t . By Proposition 4, the coefficients $u = (u_0, u_1, \dots, u_n)'$ satisfy the equation

$$Gu = m \quad (14)$$

with matrix $G_{ij} = \int_{-1}^1 x^{i+j} dx = \frac{1}{i+j+1}(1 - (-1)^{i+j+1})$ for $i, j = 0, 1, \dots, n$. This Hilbert matrix has the shape

$$G = 2 \begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & \dots \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \dots \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \dots \\ 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \dots \\ \vdots & \cdot & \cdot & \cdot & \cdot & \dots \end{pmatrix}$$

Equation (14) splits into an even part for computing u_0, u_2, \dots , and an odd part for computing u_1, u_3, \dots . The equation $G_* u_* = m_*$ for the odd coefficients $u_* = (u_1, u_3, \dots)'$ has right hand side $m_* = (m_1, m_3, \dots)' = 0$ and u_* vanishes since we approximate a measure which is symmetric around the origin. The coefficients $u^* = (u_0, u_2, \dots)'$ of even powers of x depend on the even moments $m^* = (m_0, m_2, \dots)'$ and are determined by $G^* m^* = u^*$ with matrix

$$G^* = 2 \begin{pmatrix} 1 & \frac{1}{3} & \frac{1}{5} & \frac{1}{7} & \dots \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{7} & \frac{1}{9} & \dots \\ \frac{1}{5} & \frac{1}{7} & \frac{1}{9} & \frac{1}{11} & \dots \\ \frac{1}{7} & \frac{1}{9} & \frac{1}{11} & \frac{1}{13} & \dots \\ \vdots & \cdot & \cdot & \cdot & \dots \end{pmatrix}.$$

G^* is, up to the factor 2, the Hilbert matrix studied in [9] and shown to be invertible. Note that the entries do not depend on t . This is good for numerical computations, since one has to store the inverse only once for all t .

Figure 2 shows approximations of Bernoulli convolutions for $t = 0.6$ and $t = 0.8$.

7. REMARKS ON SINGULAR MEASURES

What happens if ν is a singular measure? We tried to prove that our approximating polynomials v_n , taken as measures $\nu_n(B) = \int_B v_n(x) dx$, will converge weakly to ν , in the sense that $\nu_n(f) \rightarrow \nu(f)$ for continuous functions f . Fortunately, a referee found a gap in our proof, and an error in an attempt to repair the proof. We are very grateful for the referee's careful work! Let us briefly discuss the approximation of singular measures.

Proposition 6. (*Properties of approximating polynomials*)

Under the assumptions of Theorem 2, the v_n satisfy

- (1) $\langle v_n, q \rangle = \nu(q)$ for all polynomials q of degree $\leq n$.
- (2) $\langle v_n, q \rangle \geq 0$ for all polynomials q of degree $\leq n$ with $q(x) \geq 0$ for $x \in X$.
- (3) $\|v_n\|^2 \leq \max_{x \in X} v_n(x)$.
- (4) If ν does not admit an L^2 density then $\|v_n\| \rightarrow \infty$.

Proof. (1) was proved in Section 4, and (2) is an immediate consequence. For (3), let $M = \max_{x \in X} v_n(x)$. Then $q = M - v_n$ is a positive function. (2) implies $\langle v_n, M - v_n \rangle \geq 0$ since $\langle 1, v_n \rangle = 1$.

For (4), we note that $\|v_n\|$ is an increasing sequence since v_n is the orthogonal projection of v_{n+1} onto P_n . If the $\|v_n\|$ are bounded, they form a weakly relatively compact set in L^2 , and a subsequence of v_n weakly converges to some $v \in L^2$. Since (1) implies $\langle v, q \rangle = \nu(q)$ for all polynomials q which are dense in L^2 as well as in $C(X)$, this v must be an L^2 density of ν . So for singular ν the sequence $\|v_n\|$ tends to infinity. This kind of argument is known as uniform boundedness principle, cf. [16, Corollary 1.7.7]. \square

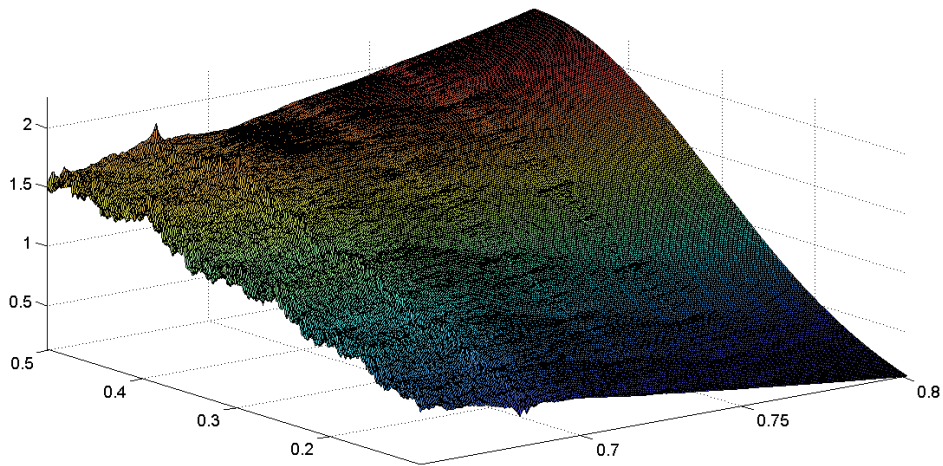


FIGURE 6. Polynomial approximation to the two-dimensional Bernoulli density for $0.65 \leq t \leq 0.8$, $0.1 \leq x \leq 0.5$. Apparent singularities at $t = 0.68$ correspond to the Pisot number $\beta = 1.4656$. For the first two Pisot numbers the singularity is less obvious.

The proposition indicates that approximation of singular measures requires unbounded and highly oscillating polynomials. For instance, (3) shows that the set $S = \{x \mid v_n(x) \geq \frac{M}{2}\}$ fulfils $M \geq \|v_n\|^2 \geq \int_S v_n^2 \geq (\frac{M}{2})^2 \cdot |S|$ and thus has Lebesgue measure $|S| \leq \frac{4}{M}$ which is very small for very large M . Such polynomials seem not to be helpful for a study of specific singular measures.

It seems appropriate, however, to study singular measures and polynomials within the context of a parametric family where most measures have an L^2 density. Due to their construction, the v_n continuously depend on the parameter for each n . We have good polynomial approximations in L^2 for all regular parameter values. For the singular values, we

can study the singularities of the two-variable function which is obtained in this way. Figure 6 shows a two-variable polynomial approximating the family of Bernoulli convolutions for $0.65 \leq t \leq 0.8$. There are three Pisot numbers $\beta = 1.3247, 1.3803, 1.4656$ where ν is known to be singular [15]. The figure shows a very smooth function: a polynomial of degree 28 in t and x . The singularities at $\beta = 1.4656$ are obvious, those at the other parameters are barely visible. A study of Bernoulli convolutions as a two-parameter L^2 function was initiated in [2] and will be continued elsewhere.

REFERENCES

- [1] David H. Bailey, Jonathan M. Borwein, Richard E. Crandall, and Michael G. Rose. Expectations on fractal sets. *Appl. Math. Comput.*, 220:695–721, 2013.
- [2] Christoph Bandt. The two-dimensional density of bernoulli convolutions. *arXiv1604.00308v1*.
- [3] Krzysztof Barański. Dimension of the graphs of the weierstrass-type functions. In Christoph Bandt, Kenneth Falconer, and Martina Zähle, editors, *Fractal Geometry and Stochastics V*, volume 70 of *Progress in Probability*, pages 77–91. Springer International Publishing, 2015.
- [4] Michael Barnsley and S. Demko. Iterated function systems and the global construction of fractals. *Proc. R. Soc. Lond.*, 399:243–275, 1985.
- [5] Bernhard Beckermann. The condition number of real Vandermonde, Krylov and positive definite Hankel matrices. *Numerische Mathematik*, 85(4):553–577, 2000.
- [6] Man-Duen Choi. Tricks or treats with the Hilbert matrix. *Amer. Math. Monthly*, 90(5):301–312, 1983.
- [7] Dean J. Driebe. The Bernoulli Map. In *Fully Chaotic Maps and Broken Time Symmetry*, volume 4 of *Nonlinear Phenomena and Complex Systems*, pages 19–43. Springer Netherlands, 1999.
- [8] Kenneth J. Falconer. *Fractal geometry : mathematical foundations and applications*. J. Wiley & sons, 3 edition, 2014.
- [9] David Hilbert. Ein Beitrag zur Theorie des Legendre’schen Polynoms. *Acta Mathematica*, 18:155–159, 1894.
- [10] John E. Hutchinson. Fractals and self-similarity. *Indiana University Mathematics Journal*, 30:713–747, 1981.
- [11] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer Berlin Heidelberg, 1976.
- [12] Pablo Shmerkin. On the Exceptional Set for Absolute Continuity of Bernoulli Convolutions. *Geometric and Functional Analysis*, 24(3):946–958, 2014.
- [13] J.A. Shohat and J.D. Tamarkin. *The problem of moments*. Mathematical surveys and monographs. American Mathematical Society, 1943.
- [14] Boris Solomyak. Notes on Bernoulli convolutions. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 1*, volume 72 of *Proceedings of Symposia in Pure Mathematics*, pages 207–230. AMS, 2004.
- [15] Boris Solomyak, Yuval Peres, and Wilhelm Schlag. Sixty years of Bernoulli convolutions. *Progress in Probability*, 46:39–65, 2000.
- [16] Terence Tao. *An epsilon of room I*. Graduate Studies in Mathematics. American Mathematical Society, 2010.

Christoph Bandt, Helena Peña
 Institute of Mathematics, University of Greifswald, Germany
 bandt@uni-greifswald.de, thaki11110@gmail.com